DOCUMENT RESUME

ED 102 235 UD 014 703

TITLE	National Assessment of Educational Progress. General Information Yearbook.
INSTITUTION	Education Commission of the States, Denver, Colo.
SPONS AGENCY	Carnegie Corp. of New York, N.Y.; Ford Foundation, New York, N.Y.; National Center for Educational Statistics (DHEW/OE), Washington, D.C.
REPORT NO	R-03-04-GIY
PUB DATE	Dec 74
NOTE	56p.
AVAILABLE FROM	Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 ($1.20)
EDRS PRICE	MF-$0.76 HC-$3.32 PLUS POSTAGE
DESCRIPTORS	*Academic Achievement; Age Differences; *Annual Reports; Geographic Distribution; Longitudinal Studies; Mathematics Education; Music Education; *National Surveys; Research Methodology; Science Education; Sex Differences; Social Studies; *Statistical Analysis; *Yearbooks

ABSTRACT

This, the second General Information Yearbook, is designed to assist the reader in understanding Assessment findings. Report 03/04-GIY relates especially to the learning areas assessed during the third and fourth years of the project: social studies, music, mathematics, and science. It presents an overview of National Assessment, describing its purpose, history, and methodology, and then goes on to devote a chapter to each of the major steps in the intricate research/dissemination process. The chapters are organized to tell how the assessment is carried out and subsequently how its data are reported and used. National Assessment's (NAEP) principal objective is to assess the changes over time in educational achievement of four age groups in the U.S. In addition to the national results for each age group, results are reported within each age level for certain major subpopulations, such as geographic region and sex. To do so, NAEP had developed two baseline measures: (1) an estimate of the percentage of persons in each of the populations who gave a certain response to a specific exercise, and (2) an estimate of the difference between the performance of a subpopulation and the performance of the total age population. Measures of change will reflect comparisons of these statistics over time. (Author/JM)

ED102235

BEST COPY AVAILABLE

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

GENERAL INFORMATION YEARBOOK

Report No. 03/04-GIY

December 1974

LD 014 703

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. [illegible]

BEST COPY AVAILABLE

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

J. Stanley Ahmann
Director

George H. Johnson
Associate Director

This publication was prepared and produced pursuant to agreements with the National Center for Educational Statistics of the U.S. Office of Education with additional funds from the Carnegie Corporation of New York and the Ford Foundation's Fund for the Advancement of Education. The statements and views expressed herein do not necessarily reflect the position, and policy of the U.S. Office of Education or other grantors but are solely the responsibility of the National Assessment of Educational Progress, a project of the Education Commission of the States.

BEST COPY AVAILABLE

TABLE OF CONTENTS

LIST OF EXHIBITS v

PREFACE vii

CHAPTER 1 Overview of National Assessment 1
History and Purpose 1
Goals of the Assessment 2
Methodology 3
The Assessment Cycle 4

CHAPTER 2 Sampling 5
Design Specifications 6
In-School Sample Design 6
Out-of-School Sample Design 7
Sample Sizes 8
Response Experience 8

CHAPTER 3 Objectives and Exercise Development 13
Development and Review of Objectives 13
Exercise Preparation and Local Tryouts 13
Exercise Review and Revision 14
Field Testing, Exercise Scoring and Review of the Results 14
Final Exercise Review and Selection 15

CHAPTER 4 Preparation of Materials and Administration of the Assessment 17
Preparation of Materials 17
Administration of the Assessment 18

CHAPTER 5 Categorizing Responses and Preparing the Data Files 21
Receipt Control 21
Ongoing Hand Scoring 21
Programming and Machine Processing 22
Special Scoring 23
Special Projects 23

CHAPTER 6 Data Processing 25
Computer Hardware and Languages 25
Preparation for an Assessment Year 25
Organizing the Data 26
Availability of Data 27

CHAPTER 7 Data Analysis
Computation of F-Values
DF Values
Subgroup Differences
Summary Benefits
Estimation of Change
Limitations of the Data

CHAPTER 8 Reporting and Disseminating the Findings

CHAPTER 9 Adaptations and Utilization of the Assessment Model in
Consultation
Technical Assistance
Workshops and Seminars

APPENDIX A Definitions of National Assessment Reporting Populations
National Assessment Groups

APPENDIX B Base-year Population Distributions of National Assessment Reporting Groups

APPENDIX C National Assessment Policy Committee

APPENDIX D National Assessment Analysis Advisory Committee

LIST OF EXHIBITS

EXHIBIT 2-1. National Assessment Reporting Sample 6

EXHIBIT 2-2. National Assessment Data Map Sheet 7

EXHIBIT 2-3. Tar Yield (Mean) and Receipt Brackets (Reported Sample Size 9

EXHIBIT 2-4. Tar Yield (Median) and Methodology (Reported Sample Size 9

EXHIBIT 3-1. Average Quantity of Responses by Group and Type of Administration at Each Age Level: Tar Yield (Mean) and Receipt Brackets 10

EXHIBIT 3-2. Average Quantity of Responses by Group and Type of Administration at Each Age Level: Tar Yield (Median) and Methodology 10

EXHIBIT 7-1. Mean Differences Among the National Percentage of Receipts and Reported Sample Weight 37

EXHIBIT 2-1. National Assessments Populations, Survey Periods and Age Groups 40

EXHIBIT 3-2. Distributions of National Assessment Regional Subpopulations 41

EXHIBIT 3-3. Occupational Categories 42

EXHIBIT 3-4. National Assessment Size and Type of Community (MSA) Reporting Categories 43

EXHIBIT 3-5. Distributions of Respondents by Size and Type of Community (MSA) and Age Level 43

EXHIBIT B-1. Annualized Population Distributions of National Assessment Reporting Groups (All Exam Males) 44

EXHIBIT B-2. Annualized Population Distributions of National Assessment Reporting Groups (All Exam Males) 45

EXHIBIT B-3. Annualized Population Distributions of National Assessment Reporting Groups (All 21 Exam Males) 46

EXHIBIT B-4. Annualized Population Distributions of National Assessment Reporting Groups (Young Adults) 46

[illegible]

[illegible] is the second biennial conference from [illegible] conference to the Electronic Assessment of Mitochondrial Myopathy Disorder. It is designed to be a [illegible] for understanding [illegible] biology. Research [illegible] at [illegible] is the following three sessions during the third and fourth years of the project: social studies, [illegible] and evidence of personnel on [illegible] of [illegible] [illegible], according to [illegible] [illegible] and methodology, and then goes on to [illegible] a report on each of the major areas of the [illegible] research [illegible] process.

The [illegible] of the [illegible] [illegible] Yudkowsky are [illegible] to aid [illegible] the [illegible] [illegible] and [illegible] [illegible] [illegible]. It [illegible] are [illegible] and [illegible]. The [illegible] provides a [illegible] [illegible] [illegible] of [illegible], [illegible], [illegible] and [illegible] and [illegible] [illegible]. [illegible] [illegible] [illegible]

BEST COPY AVAILABLE

CHAPTER 2

SAMPLING[1]

National Assessment's (NAEP) principal objective is to assess the changes over time in educational achievement of four age groups in the United States. In addition to the national results for each age group, results are reported within each age level for certain major subpopulations, such as geographic region and sex. Current reporting populations and subpopulations are shown in Exhibit 2-1; the definitions for each are presented in Appendix A.

EXHIBIT 2-1. National Assessment Reporting Groups

Classification	Subgroup
Age level	9-year-olds 13-year-olds 17-year-olds Young adults (26-35 years)
Sex	Male Female
Geographic region	Northeast Southeast Central West
Level of parental education	No high school Some high school Graduated high school Post high school
Size and type of community (STOC)	Low metropolitan High metropolitan Extreme rural Main big city Urban fringe Medium city Small places
Color	White Black

There are two ways to meet this objective: conduct a census of all members of the four age populations or utilize probability samples of the four populations. A complete census has the advantages of providing great precision and of allowing very detailed reporting. However, the four populations are very large (about 37 million people altogether); finding and assessing that many people would involve enormous expenditures of time and resources for data collection, scoring and analysis. Furthermore, it would be extremely difficult to adequately train and supervise the large field staff necessary to assess all members of the population.

Since National Assessment only collects data about a limited number of subpopulations and not individual students, schools, school districts or even states, relatively small probability samples allow estimation of educational performance in these subpopulations with the required degree of precision. Further, by limiting the scale of data collection activities, it is possible to use a small, highly trained professional field staff to collect data of higher quality than would be possible with a census. Quantity of data is traded for much higher quality at reasonable cost.

[1] This chapter provides a brief overview of National Assessment sample designs for assessment Years 02–04. For more detailed documentation, see J.R. Chromy et al., *The National Assessment Approach to Sampling* (Denver, Colorado: National Assessment of Educational Progress, 1974). For an introductory treatment aimed at state and local assessment problems, see R.M. Jaeger, *A Primer on Sampling for Statewide Assessment* (Princeton, N.J.: Educational Testing Service, Center for Statewide Educational Assessment, 1973).

BEST COPY AVAILABLE

The development of sample designs was based upon a long series of interactions between what the Assessment ought optimally to accomplish and what it could accomplish, given the available finite resources. Policy, data collection, analysis, reporting, cost and time considerations determined what types of sample designs were possible. The sample designs, in turn, put major constraints on the options available for data collection, analysis and reporting. For example, the limits placed on the number of administrations in a school and the size of group administrations (12 students) are a compromise between cost efficiency and practical feasibility considerations.

Design Specifications

The target populations in the Year 03 and 04 assessments included 9, 13 and 17-year-olds enrolled in public or private school, 17-year-olds who either left school before graduating or graduated early and young adults 26 to 35 years old. Age-eligible persons who were non-English speaking, institutionalized or handicapped (physically, mentally or emotionally) in such a way that they could not respond to the exercises as administered, were excluded.

At each age level the sample for each package of exercises was designed to meet the following specifications:

1. Adequate representation of the subpopulations to allow estimation of the desired proportions with an acceptable level of precision.

2. Representation by at least one sample point for each of the states and the District of Columbia. The design was not, however, to provide for making comparisons among states, school districts, schools, teachers or individual students.

3. Facilitation of field-operating procedures.

NAEP's policy that administration of materials was not to take more than one class period of a student's time, that the demands on school personnel were to be minimized (one package per student) and that the number of students assessed in any one school was to be limited to 12 also placed constraints on the sample design. Among other important sampling considerations were (1) that no more than about one half of the group-administered packages were to be administered in any one school, (2) that out-of-school 17-year-olds and young adults were to be allowed to take up to four packages and (3) that the respondents taking each package were to be a probability subsample of the total sample for the age group.

This design provided for simple, precise estimates of population proportions. Since results were to be reported by individual exercise, these estimates were to be relatively inexpensive to compute.

To meet these specifications, two three-stage, deeply stratified cluster designs with extra sampling of certain strata were developed. One design was for the survey of 9, 13 and 17-year-olds enrolled in school. The out-of-school design was developed to assess young adults aged 26 to 35 and 17-year-olds not enrolled in school.

In-School Sample Design

Stratification

In a stratified design the population is divided into two or more groups, or strata. Samples are then drawn from each stratum rather than from the population as a whole. The two major reasons for stratification are to insure representation of specified subpopulations and to achieve more precise estimates. National Assessment samples were stratified by geographic region, state, size of community and socioeconomic level as shown in Exhibit 2-2. Socioeconomic stratification was performed at the school level in the two largest size of community (SOC) strata and at

BEST COPY AVAILABLE

the primary sampling unit level in the remaining two size of community strata.

EXHIBIT 2-2. National Assessment Sampling Strata

Classification	Strata
Region	Northeast
	Southeast
	Central
	West
Size of community (SOC)	Counties containing cities with 1970 populations of 180,000 or greater
	Counties in the same Standard Metropolitan Statistical Area (SMSA) as the above counties
	Counties not in the above strata that are part of an SMSA or contain a city with a 1970 population of 25,000 or greater
	All other counties
Socioeconomic level	High—schools in the first two SOC strata
	Low—counties in the last two SOC strata

Multi-Stage Sampling

A multi-stage design involves sampling in successive steps or stages in order to control sampling and data collection costs.

In National Assessment's in-school sample design, the first stages—or primary sampling units (PSUs)—consisted of counties or groups of contiguous counties. A sampling frame of PSUs was constructed with United States Census data on the number of persons by age in each PSU. The PSUs were stratified by region, and within region by state, size of community and, for the two smaller size of community strata, by socioeconomic level. From this stratified list of PSUs, a probability sample of 116 PSUs was independently drawn each year.

At the second stage, a list was made of all public and private schools in each selected PSU. The listing included the estimated number of 9, 13 and 17-year-olds enrolled in each school. A secondary sample of schools was then drawn for each sample PSU. Schools in the two larger size of community strata were stratified by socioeconomic level prior to selecting the secondary sample.

In each selected school, the third-stage sampling units consisted of eligible students enrolled in the school. Every eligible student was listed. A random sample of students was then drawn and randomly assigned one of the assessment packages scheduled for that school.

To ensure adequate sample sizes for each of the various reporting groups listed in Exhibit 2-1, it was necessary to increase the sample size of the population at the lower end of the socioeconomic scale. This "oversampling" was performed at the first stage in the two smaller size of community strata and at the second stage in the two remaining size of community strata.

Out-of-School Sample Design

The out-of-school sample design is similar to the in-school design. As with the in-school design, primary sampling units consisted of counties or groups of contiguous counties stratified by region, state, size of community and socioeconomic level. One hundred four primary sampling units were drawn and used for both Year 03 and Year 04 (one half of them were used in Year 02). Each year, a new secondary sample of small land areas (segments) was drawn in each PSU. All of the housing units in each segment were listed and a sample of the housing units was screened for eligible out-of-school 17-year-olds and adults.

The Supplementary Frame

In order to increase the sample size of 17-year-olds not enrolled in school, a probability subsample of one half of the schools in the 17-year-old assessment was drawn. Each school was asked to provide lists of persons

BEST COPY AVAILABLE

who had either graduated early or dropped out of school. A subsample of dropouts and early graduates was drawn.

Sample Size:

The size of a sample required for estimating baseline performance and change over time is related to the type of administration, the minimum change to be detected with a given degree of confidence, the desired power, the sample design, operational procedures, time and cost. In accounting for these, NAEP set the planned sample size for the in-school samples at 12 individuals per group-administration session and the planned national sample sizes at 2,592 for group-administered packages and 2,160 for individually administered packages.

One problem recognized in planning the sample sizes for the in-school assessment was the reduction in sample size due to absenteeism at the time of assessment. Therefore, to assure that the desired sample sizes were achieved, a random sample of 16 students was selected for each group administration. The first 12 students were assigned to the assessment session. The remaining 4, designated as alternates, were used to replace any of the 12 students who were absent at the time scheduled for the package administration. For each individual administration, 2 students were randomly selected with the second serving as an alternate for the first. Special selection procedures were adopted for the in-school sample to accommodate schools enrolling less than the required number of students for one group administration. The allocations to these schools were in addition to the planned national sample sizes given above.

Response Experience

The response experiences for Year 03 and Year 04 are summarized in Exhibits 2-3 through 2-6. The observed sample sizes per package are presented in Exhibits 2-3 and 2-4. Exhibits 2-5 and 2-6 present the average size per package for each of the reporting groups. The planned sample sizes were 2,592 respondents per package for in-school group administrations, 2,160 per package for in-school individualized administrations and 2,000 respondents for each package in the out-of-school sample. The exhibits reveal that the observed sample sizes were very close to the projected sizes.

EXHIBIT 3-3. Year 03 (Music and Social Studies) Observed Sample Size

Package Number	Age Level			
	9	13	17	Adult
1	2,747	2,661	2,654	2,393
2	2,712	2,665	2,679	2,438
3	2,733	2,688	2,554	2,350
4	2,662	2,664	2,707	2,395
5	2,746*	2,647	2,672	2,367
6	2,721	2,741	2,659	2,398
7	2,741	2,655	2,665	2,356
8		2,760*	2,670	2,397
9		2,747	2,722	
10		2,747	2,649*	
11			2,806	
12			2,697	
Packages administered:	17,781	21,569	31,685	18,766
Respondents:	17,781	21,569	28,812	4,822

Total number of packages administered at all ages: 93,731
Total number of respondents at all ages: 76,984*

*Sample sizes below line are for individually administered packages.
†Adults and out-of-school 17-year-olds were allowed to take more than one package.

EXHIBIT 3-4. Year 04 (Science and Mathematics) Observed Sample Size

Package Number	Age Level			
	9	13	17	Adult
1	2,666	2,931	2,671	2,997
2	2,670	2,698	2,723	2,987
3	2,628	2,674	2,756	2,999
4	2,646	2,621	2,723	2,898
5	2,670	2,684	2,791	2,838
6	2,674	2,512	2,729	2,876
7	2,684	2,689	2,672	2,858
8	2,758*	2,570	2,682	2,829
9	2,283	2,573	2,638	
10	2,224	2,569*	2,633	
11		2,254	2,682	
12		2,237	2,587*	
13			2,582	
14			2,498	
Packages administered:	25,403	36,247	37,890	16,358
Respondents:	25,403	36,247	33,577	4,233

Total number of packages administered at all ages: 129,830
Total number of respondents at all ages: 93,630

*Sample sizes below line are for individually administered packages.
†Adults and out-of-school 17-year-olds were allowed to take more than one package.

EXHIBIT 14. Average Number of Responses by Group and Type of Administration at Each Age Level: Year 01 (Nurse and Social Worker)

	Age Level						
	03				07		4AHR
	Group	Individual	Group	Individual	Group	Individual	
National	1,784	2,997	2,472	2,246	2,474	2,858	2,581
Region							
Northeast	674	774	671	729	688	685	674
Southeast	696	810	679	742	765	698	742
Central	668	776	641	683	667	672	672
West	669	678	661	747	669	672	673
Sex							
Male	1,734	1,980	1,247	1,146	1,256	1,289	1,868
Female	1,848	2,247	1,235	1,188	1,865	1,284	1,295
Color							
White	7,342	1,746	2,136	1,832	2,177	2,830	1,928
Black	818	850	342	365	330	311	380
Other	384	358	193	147	167	177	177
Parental education							
No high school	237	234	196	248	264	288	720
Some high school	388	422	263	233	272	288	358
Graduated high school	647	952	829	787	816	761	842
Post high school	876	764	1,825	878	1,072	1,862	895
Unknown	862	725	365	313	489	234	1
Size and type of community (SESAC)							
Low metro	211	196	248	227	312	256	232
Extreme rural	276	211	253	234	228	264	236
Small place	677	788	858	772	896	826	786
Medium city	858	867	863	276	838	808	872
Main big city	263	282	244	256	285	234	288
Urban fringe	271	255	228	243	268	245	268
High metro	212	243	267	286	289	229	218

[illegible]

[illegible]

	[illegible]						
	[illegible]						
	Average	Male/female	Average	Male/female	Average	Male/female	
National	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Region							
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Sex							
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Color							
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Parental education							
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Size and type of community (STRATA)							
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

[illegible]

The development of objectives and exercises include the following:

1. development and review of objectives;
2. exercise preparations and local issues;
3. exercise review and analysis;
4. field training, exercise scoring and review of the results; and
5. final exercise review and selection.

Development and Review of Objectives

Effective objectives must distinguish the new area which exists, include emphasis on external input for previously assessed areas, exercises the final criteria had to satisfy a wide variety of participants' perceptions and allowances in each training area. Acceptance and approval of each set of objectives was dependent on all simultaneous goals and also showed those concerned criteria. Wisconsin Statewide Assessment (WSAT) has maximum consistency for known conditions and approach of materials, eliminate input from each to model experience, and those all aspects of the counties, participating lists of questions of participating groups and participations of both states and locations.

In addition, the assessment selected subjects, higher perceptions, and allowances representing different perceptions within the descriptions of federal organizations and projects exercises, communications.

Objectives development and review is specific plan for [illegible] the description to the [illegible] of the [illegible] to [illegible] of the [illegible] and [illegible] with the training area as discussion and local assessment, and secondary levels of public, private and parochial schools provide unique local issues, implications, [illegible] concerns, among [illegible].

Explicit selection requirements for [illegible] of [illegible] participant representatives in local, school, and levels of participations. Review [illegible] organizations and different participations.

Objectives development must [illegible] and also to accomplish the each training area. Consideration and consultations [illegible] of local areas [illegible] within each area to accurately [illegible] include, implying and of educational objectives and provide [illegible] organizations. Then [illegible] selection and [illegible] each [illegible] to a select [illegible] development and [illegible] [illegible] and which [illegible] to the [illegible] of objectives. Further review, including a review by local [illegible], [illegible] a [illegible], local [illegible] of these the objects of [illegible] and assessment, the final [illegible] and [illegible] and assessment by an [illegible] [illegible] of subject, [illegible] [illegible], the [illegible] [illegible] objectives, [illegible] [illegible] to the [illegible] assessment.

Exercise Preparations and Local Issues

Based on the objectives the each training area, [illegible] developed specifications for the [illegible] [illegible] and quality of [illegible] to be developed. Special emphasis was given to [illegible] considerations, [illegible] [illegible], [illegible], [illegible] and [illegible] of [illegible] [illegible] and [illegible] level. Because [illegible] [illegible] to be [illegible] [illegible] [illegible] and [illegible] to the [illegible] [illegible] [illegible] [illegible] [illegible], [illegible] [illegible], [illegible] [illegible] [illegible] [illegible]

[illegible]

[illegible]

[illegible]

BEST COPY AVAILABLE

CHAPTER 5

CATEGORIZING RESPONSES AND PREPARING THE DATA FILES

The data that National Assessment (NAEP) collects in the field must be coded into a form suitable for computer processing and analysis. Measurement Research Center (MRC) of Iowa City, Iowa, was the contractor in charge of receiving, categorizing and machine processing the data. Responses to open-ended and performance exercises were categorized and coded into machine-readable form by a special staff of scorers. Responses to multiple-choice exercises were read directly by optical scanning machines. These machines transferred the information to computer data tapes which provided the basis for data analyses performed by the NAEP staff.

Receipt Control

When booklets were printed and shipped to the field staff, each one-week unit of work was packed in a fiberboard box. Once the packages were administered, they were repacked in the same box and returned to MRC for processing and scoring. From the beginning to the end of the MRC processing, that unit of work was kept intact.

Upon receipt from the field, each box was unpacked by the receipt control staff and checked to ensure that:

1. All forms and worksheets were properly and accurately completed.
2. All materials (packages, tapes, work samples, etc.) were properly and uniquely identified.
3. All materials were accounted for.
4. Personal data and documentation for each respondent were complete and consistent.
5. Machine-readable responses were gridded according to instructions and were dark enough to be scorable.
6. Administrator-recorded responses were legible and scorable.
7. Branching exercises were being correctly administered.

MRC corrected errors and inconsistencies which could be solved with the information at hand. Problems for which there was insufficient information were referred to the field staff for solution. In all cases, NAEP staff was notified of the problems and their recommended solutions. After the receipt control checks were completed, each box was repacked and given to the MRC hand-scoring staff.

Ongoing Hand Scoring

In each assessment year, MRC employed a special staff to hand score open-ended exercises. Scorers were responsible for categorizing responses according to predetermined guidelines and coding the category of the response into ovals that could be read by the optical scanning machine.

Hand-scoring tasks ranged from simple tallies to rather high-level judgments. Hand scorers were required to have some academic background and, preferably, teaching experience in the subject they were scoring.

BEST COPY AVAILABLE

When the initial data (age 13) began to arrive from the field, the scoring staff was assembled for training. A training session took place prior to the scoring of each age level. The first training session was the most complex, since scorers had to become familiar with NAEP, the objectives related to the particular learning area and general scoring and coding procedures.

The MRC director of scoring conducted the training sessions in cooperation with NAEP staff. Once the general orientation was accomplished, training began on specific exercises and their scoring guides. The guides were initially developed by the exercise developer, reviewed and refined by subject-matter and lay consultants and refined further by NAEP staff after tryouts. Acceptable and unacceptable responses were defined through prose descriptions and sample responses. Response categories that were of interest to NAEP staff and consultants were also defined.

Scorers were trained in the use of the guides by scoring sample responses, which were either taken from tryout material or the arriving assessment data. Scorers initially worked as a group and discussed the appropriate categorization of each example response. They then worked individually on another set of responses. Discrepancies were resolved and explained. Once the group felt comfortable using the guides, they started categorizing the actual data. The director of hand scoring checked the scoring for the first 20—24 papers for each exercise for each scorer. Incorrect categorizations were explained and careless errors were pointed out. Subsequently, spot checks were made on the work of each scorer and retraining was conducted as needed.

Based on the final data some scoring guides needed modification. Generally, this involved some redefinition of scoring categories. All responses scored before a scoring guide change were, of course, rescored.

Scorers were approved by NAEP prior to training and could be removed at any time if their work was found unsatisfactory. Quality-assurance procedures were developed both by MRC and NAEP to ensure the consistency of hand scoring. Members of the NAEP staff made several trips to MRC to observe and assist in training sessions and to check that the hand scoring was consistent with the scoring guides furnished by the project.

After all the data at a given age level was scored, MRC prepared a report which included the final form of the scoring guides, examples of responses for scoring guide categories and a summary of the problems encountered and their solutions.

Programming and Machine Processing

Once the open-ended responses were categorized, the packages were repacked and sent on for machine processing. While the exercises were undergoing hand scoring, all of the procedures and programs to be used in the machine processing were being tested on materials developed by MRC and NAEP staff. The optical scanning device was programmed to read the response data on the exercise pages and produce data tapes which could be used by computers.

Once the packages arrived at the MRC processing department, the pages were separated so they could be passed through the scanning device. Each page contained preprinted, machine-readable information such as age, package, exercise and page number. On pages with multiple-choice exercises the respondents had made pencil marks in ovals next to the answers they chose. For other exercises the hand-scoring staff had marked codings in preprinted oval positions on the page. The scanning device, through the use of reflected light, distinguished the position of each pencil-marked oval. The scanner assigned numerical values, specified in computer programs, to the various marks according to their position on the page. The coded information was transferred onto magnetic tapes which later were used to construct the NAEP respondent data files.

BEST COPY AVAILABLE

MRC's computer checked and edited the contents of the magnetic tapes for package counts, internal exercise consistency and response ranges before they were sent to National Assessment. Tapes were also provided to RTI, the sampling contractor, for calculating weights.[1] After scanning and editing, exercises which required special scoring were isolated. The remaining pages were then sent to a climate-controlled warehouse for permanent storage. It is a policy of NAEP to save all source documents.

Special Scoring

Occasionally, the scoring of an exercise requires special scorer qualifications and training.

Music exercises involving the evaluation of singing and playing an instrument, a mathematics exercise about computer programming and several social studies and science exercises were scored professionally by subject-matter experts. The mathematics exercise was scored by bona fide computer programmers who read only that one exercise. Professional scorers were trained to score all of the musical performance exercises. Judgments were made based on the criteria established in scoring guides. For example, music scorers listened to tape recordings of respondents singing or playing instruments and then scored the renditions for accuracy of pitch and rhythm according to specified criteria.

Special scoring often involved separate scheduling procedures. The ongoing hand scorers left the special exercises blank. These exercises went through the normal machine processing procedures and then were pulled from the packages and held until the time of the special scoring session. These sessions were held after nearly all the regular scoring and processing for the assessment had been completed. After scoring, the exercise pages were rescanned and data tapes were created. The tapes were later combined to construct National Assessment respondent data files.

Special Projects

Each year, NAEP and MRC staffs also carried out various tryouts, feasibility studies and validity studies, which are an integral part of NAEP's development, quality-assurance and feedback program. These special studies varied depending on the requirements of the learning area being assessed.

[1] For a discussion of weighting procedures, see Chapter 7, Computation of P-Values.

BEST COPY AVAILABLE

CHAPTER 6

DATA PROCESSING

The data gathered by Research Triangle Institute (RTI) and Measurement Research Center (MRC) were combined, organized and analyzed by National Assessment (NAEP). Specifications for analyses to be performed were supplied by NAEP's Research and Analysis Department staff. NAEP's Data Processing Department staff is responsible for providing fast, easy access to the data on a daily basis. They also engage in long-range planning and developmental work and provide general technical support to all the project departments and their advisory committees.

Computer Hardware and Languages

National Assessment uses an IBM 360 Model 20 computer located in Denver, an IBM 360 Model 91 computer located at the Princeton University Computer Center and a high-speed communications link between the two computers. Card decks are read into an IBM 2560 card reader in Denver and transmitted to the Model 91 at Princeton for execution. Printouts are transmitted from Princeton to an IBM 1403 Printer in Denver.

Because of the large amount of data collected by NAEP, the National Military Command Information Processing System (NIPS—a Data Base/Data Management System) is used to process NAEP data. It provides a convenient method for solving the data-handling problems associated with high-volume and large-file processing requirements by using executive programs which simplify common operations such as data definition, data storage, maintenance and retrieval; the amount of human effort required to code, test and debug is reduced.

In addition to NIPS, various general-purpose programming languages, chiefly PL/I and FORTRAN are used. These languages are well suited for statistical computations as well as data manipulations on the various files.

Preparation for an Assessment Year

As the data were being collected in the field, files that would hold the data and the procedures that would be used to manipulate them were being designed at National Assessment. MRC and NAEP discussed specifications for the format in which the data would be prepared by MRC and the edit checks that MRC would perform to insure correct, accurate data. To insure compatibility, formats were also determined for data from special scoring.

Approximately three months before the first MRC-produced data tape was scheduled to be completed, file design and coding began for the respondent file. The magnetic tapes containing response data for each age level were received from MRC at six-to-eight-week intervals. The data was loaded onto the NAEP respondent files as it was received. Approximately four weeks after the MRC tape for an age level was received, a tape containing a weighting factor for each respondent was received from RTI, and the weights were merged with the respondent file.[1]

[1] Refer to Chapter 7, Computation of P-Values, for a discussion of weighting.

BEST COPY AVAILABLE

Organizing the Data

A variety of files used to check the accuracy of the data and implement various steps in the analyses of the data were developed.

Exercise Documentation File

The exercise documentation file contained exercise-level information for each exercise, keyed by learning area, exercise number, overlap and part. The file included information about the parts of each exercise, such as a short description of the exercise part, the correct foil, the response range, specifications for combinations of parts and analysis specifications. The file also contained specifications needed by the retrieval program used to create auxiliary files.

Respondent File

The respondent file was the main source of data used for all analyses. Four respondent files were created from the MRC and RTI tapes—one for each age group. Each file was organized by package-exercise number and contained records keyed by age, package, PSU (primary sampling unit) and individual identification number—each respondent represented by one record. After the respondent data was loaded onto the files, the weighting factors were loaded along with data obtained from the principal's questionnaire or background questionnaires.

Auxiliary Data Files

Since each analysis used only a fraction of the data from the respondent file, smaller auxiliary files were created in order to increase processing efficiency. Each auxiliary file represented one age level for one learning area. Part of the process of creating these files was to produce data fields which contained combinations of original items of data for each respondent. These additional, derived data fields were stored with the original data on the auxiliary files.

After the auxiliary files had been created, preliminary analysis programs computed percentages of individuals choosing each response for each exercise. They were checked for the validity of response codings and the potential usefulness of the combinations of certain responses. If an invalid coding of a response was discovered, MRC was requested to locate the original package in order to correct the error. Corrections of the data on the respondent file were made and revised specifications for analyses were prepared. New auxiliary files were created and final computer programs were written to calculate percentages, standard errors and other statistics.

Summary Files

The statistics calculated from the auxiliary files were stored on exercise statistic summary files. One summary file was created for each learning area and was keyed by age, package, exercise, part and response. Each record contained the statistics which had been computed for a single response. Additional coded information indicating exercises with similar characteristics was also added to this file, to be used later in the computation of summary statistics across groups of exercises.

Various summary analyses were performed on data from those exercises with similar characteristics. For example, the music performance exercises were summarized, and results for the social studies exercises dealing with political knowledge and attitudes were summarized. A variety of data display programs were used to present tables of medians, means, maxima and minima as well as graphic displays of the distributions of certain exercise statistics. Such summaries provided the basis for generalizations about national and subpopulation performances on various types of exercises.

Appendix File

For each learning area an appendix file was created to produce camera-ready printouts of response statistics for NAEP statistical reports. Descriptive text for each exercise

response was coded, keypunched and edited. This was a massive task. For example, music alone had over 17,000 different response possibilities. Exercise-level statistics were extracted from the summary files and combined with the edited text to create the appendix files.

Availability of Data

In addition to statistical and prose reports produced from the various data analyses, National Assessment also produces computer data tapes for the learning areas assessed. These tapes are created from the respondent files and contain the original response data for the exercises which have been released into the public domain in their entirety. In order to protect anonymity of respondents and political units, the tapes were designed so that it is not possible to relate individual responses to particular respondents, schools, school districts or states. However, there is complete documentation of the classification systems and the contents of the tapes so that the data may be used and reanalyzed to suit the needs of researchers and other interested individuals.

CHAPTER 7

DATA ANALYSIS

National Assessment (NAEP) estimates the percentages of individuals at each of three age levels in the nation as a whole and in certain groups who possess various knowledges, understandings, skills and attitudes.

To do so, the Assessment has developed two basic[1] statistics: (1) the p-value, an estimate of the percentage of persons in each of the populations who gave a certain response to a specific exercise, and (2) the i-p-value (difference), an estimate of the difference between the performance of a subpopulation and the performance of the total age population. Measures of change will reflect comparisons of these statistics over time.

The most frequently reported p-values are those that estimate the percentage of a total age population or a certain subpopulation that gave an acceptable response to an exercise. However, NAEP computed many p-values in addition to the percentage of acceptable responses. For multiple-choice exercises, percentages were computed for each unacceptable choice, the "I don't know" response and the total percentage of the unacceptable responses as well. Open-ended exercises generally had several categories of acceptable and unacceptable responses. Percentages were computed for each category as well as for the total number of acceptable and unacceptable responses. Not all exercises had clearly definable acceptable and unacceptable criteria; for example, attitude and opinion questions, and education exercises might list the percentages of people who go beyond certain levels of science mathematics or to most certain levels of science books or magazines.

Many exercises had several parts to respond individuals to to give more than one response each of which could be scored acceptable unacceptable. Additional questions asked students to rate the value of an exercise. For these multi-part exercises the percentage of individuals getting exactly one, two, three, etc. acceptable responses was computed as well as the p-percentage of individuals getting at least one, 2, or three, or four, etc. acceptable responses.

For some exercises, subsets of combinations of responses from two or more parts provided additional interesting information. For example, the following combinations were delineated: (1) percentages of individuals who at least on disagreement with a statement and unfavorably defended their positions, (2) percentages of people who could not read answers, but played a science game successfully, and (3) percentages of respondents at who described social problems, but could suggest progress solutions for those problems.

Extrapolation of P-Values

The percentages reported by NAEP are weighted estimations of population percentages. The procedures for developing the composite and item assessment values are subject to the sample designs. In every instance that the

[1] A detailed discussion on National Assessment's statistical computation procedures is available in Robert Lapore, Wayne Martin, Todd Rogers, Ron Mann, Diane Bowman and David Wright, A Look at the Analysis of National Assessment Data presented by J. Stanley Ahmann at Problems of Educational Measurement and Information Science ... (1972) ed. William E. Coffman (Boston: Houghton Mifflin Company, 1972), pp. 97-113.

[illegible] [illegible] [illegible] [

[illegible]

[illegible]

[illegible]

BEST COPY AVAILABLE

As state and local educational agencies develop and implement assessment plans adapting the Assessment model, technologies and procedures, the MU/A program develops materials to fulfill two purposes:

1. documentation of the plans and methodologies for implementation, including both success and problem areas, for those who plan later assessments and
2. documentation of the utility of analyzed and interpreted data from state and local assessments to show other educational agencies the usefulness of assessments based upon adaptations of various components of the NAEP model.

BEST COPY AVAILABLE

APPENDIX A

DEFINITIONS OF NATIONAL ASSESSMENT REPORTING POPULATIONS

National Assessment (NAEP) reporting populations include 9, 13 and 17-year-olds enrolled in public or private schools as well as 17-year-olds not enrolled in school and young adults aged 26 to 35. Within these age groups, results are reported by geographic region, sex, color, parental education and size and type of community.

With a few exceptions, listed below, the reporting populations include all age-eligible persons in the 50 states and the District of Columbia. Nine and 13-year-old populations are restricted to those students enrolled in public or private schools during the survey period. The 17-year-old population includes students enrolled in public or private schools during the survey period, those age-eligible persons not enrolled but identified by schools as having dropped out or graduated early or those age-eligible persons not enrolled but living in identifiable housing units in the United States. Survey periods and age ranges for the four age groups are shown in Exhibit A-1. Age-eligible persons are excluded if they are:

1. non-English speaking,
2. institutionalized and
3. physically, emotionally, or mentally handicapped in such a way that they cannot respond to the exercises as administered.

EXHIBIT A-1. National Assessment Populations, Survey Periods and Age Ranges

Age Level	Year 03 Survey Period*	Year 03 Birth-Date Range*
9	January – February 1972	1/1/62 – 12/31/62
13	October – December 1971	1/1/58 – 12/31/58
17 in-school	March – April 1972	10/1/54 – 9/30/55
17 not enrolled in school during January 1972	January – July 1972	10/1/54 – 9/30/55
17 not enrolled in school during March 1971	January – July 1972	10/1/53 – 9/30/54
Young adults (Year 03)	January – July 1972	4/1/36 – 3/31/46
Young adults (Year 04)	October 1972 – May 1973	1/1/37 – 12/31/46

**For Year 04 survey periods and birth-date ranges of 9, 13 and 17-year-olds, add one year to the Year 03 dates.*

BEST COPY AVAILABLE

National Assessment Groups

Geographic Regions

Results are reported by the four geographic regions defined by the Office of Business Economics, Department of Commerce. The states in each region are shown in Exhibit A-2.

EXHIBIT A-2. Definitions of National Assessment Regional Subpopulations*

Northeast	Southeast
Delaware	Arkansas
Connecticut	Florida
Maine	Virginia
New Hampshire	West Virginia
Rhode Island	Alabama
Vermont	Georgia
District of Columbia	Kentucky
Maryland	Louisiana
Massachusetts	Mississippi
New Jersey	North Carolina
Pennsylvania	South Carolina
New York	Tennessee
Central	**West**
Iowa	Alaska
Kansas	Hawaii
Nebraska	Idaho
North Dakota	Montana
South Dakota	Nevada
Minnesota	Wyoming
Missouri	Arizona
Illinois	Oregon
Indiana	Utah
Michigan	Colorado
Wisconsin	New Mexico
Ohio	Oklahoma
	California
	Texas
	Washington

**These regional subpopulation definitions are the same as those used by the Office of Business Economics, Department of Commerce.*

Sex

Results are reported separately for males and females at all age levels.

Color

Results are reported for Whites and Blacks. While respondents were classified into several other ethnic groups, including Puerto Ricans, Mexican-American and other or unclassified, the actual sample sizes realized were insufficient for reporting purposes.

Parental Education

Parental education refers to the highest level of education level reported by the respondent for either parent.

No high school. Neither parent has any formal education beyond the eighth grade.

Some high school. At least one parent has some formal education beyond the eighth grade, but neither parent has graduated from high school.

Graduated from high school. At least one parent has graduated from high school, but neither parent has any formal education beyond high school.

Post high school. At least one parent has some formal education beyond high school including any business, professional or trade school training as well as college or university training.

Size and Type of Community (STOC)

The seven size and type of community (STOC) reporting categories are comprised of three "extreme" types of community (TOC) and four "residual" sizes of community (SOC). Each TOC category includes approximately 10% of the respondents at each age level; the remaining respondents are classified according to one of the SOC classifications.

Briefly, the three TOC categories are: (1) city areas where a high proportion of the adult population is either not regularly employed or

BEST COPY AVAILABLE

on welfare and a low proportion is employed in professional or managerial positions; (2) rural areas where a high proportion of adults are farm workers and a low proportion are professional, managerial or factory workers; and (3) near-city and city areas where a high proportion of adults are employed in professional or managerial positions and a low proportion are factory or farm workers, not regularly employed or on welfare. Respondents are placed in one of these categories if the occupational profile and location of the school or, in the case of the out-of-school sample, segment satisfy the extreme TOC definitions.

The remaining respondents at each age level are classified according to the size of community in which the school or segment is located. The occupational profile is based on the employment categories summarized in Exhibit A-3.

For the in-school sample at each age and the supplementary sample at age 17, the school principal of each selected school provided estimates of the percentage of students whose parents fit into each occupational category.

Other occupational data for the out-of-school sample was obtained from census data in Year 03 and from the respondents themselves in Year 04. The definitions used to classify respondents by STOC are presented in Exhibit A-4. The occupational index is computed using the occupational categories summarized in Exhibit A-3. The distribution of respondents by STOC category for Year 03 is presented in Exhibit A-5.

EXHIBIT A-3. Occupational Categories

Categories	Code
Professional or managerial personnel	A
Sales, clerical, technical or skilled workers	B
Factory or other blue collar workers	C
Farm workers	D
Not regularly employed	E
On welfare	F

BEST COPY AVAILABLE

EXHIBIT A-4. National Assessment Size and Type of Community (STOC) Reporting Categories

Reporting Category	Occupational Index*	Description
Low metro	E+F-A	Sample schools or segments in a city or metropolitan area of a city with a population greater than 150,000 and in the 90 - 99th percentile of the low metro index
Extreme rural	[6xE'+2A]	Sample schools or segments in communities with a population less than 8,000 and in the 90 - 99th percentile of the extreme rural index
Small place		Sample schools or segments in a community with a population less than 25,000 and not classified as extreme rural
Medium city		Sample schools or segments in a city with a population between 25,000 and 200,000 and not classified as low metro or high metro
Main big city		Sample schools or segments within the city limits of a city with a population greater than 200,000 and not classified as high metro or low metro
Urban fringe		Sample schools or segments in the metropolitan area of a big city but outside the city limits and not classified as low metro, extreme rural or high metro
High metro	A+(C'+[D+E+F)	Sample schools or segments in a city or metropolitan area of a city with a population greater than 150,000 and in the 90 - 99th percentile on the high metro index

*See Exhibit A-1

EXHIBIT A-5. Distribution of Respondents by Size and Type of Community (STOC) and Age Level

STOC Category	Age Level					
	9	13	17 (HS)*	17 (OSY)	17 (All)	Adult
1. Low metro	10.2%	9.9%	11.3%	14.0%	11.7%	9.9%
2. Extreme rural	10.0	9.6	11.5	9.8	11.6	10.2
3. Small place	32.8	33.7	33.0	33.4	33.1	31.8
4. Medium city	16.8	16.6	15.9	19.9	16.3	13.7
5. Main big city	14.4	11.2	9.3	6.4	9.0	8.0
6. Urban fringe	10.0	8.4	5.1	7.7	5.4	16.4
7. High metro	16.1	10.7	13.6	8.9	13.0	10.0

*In school
Out-of-school

BEST COPY AVAILABLE

APPENDIX B

ESTIMATED POPULATION DISTRIBUTIONS OF NATIONAL ASSESSMENT REPORTING GROUPS

Based on Year 03 data, Exhibits B-1 through B-4 contain the percentage of the population at each age level estimated to be in each two-way combination of the reporting groups defined in Appendix A. These data provide a more complete description of the composition of the population at each age level in terms of National Assessment's (NAEP) reporting variables. Each entry is an estimate of the percent of a reporting group identified by a row heading found in a second reporting group identified by a column heading. For example, in each exhibit, the rows corresponding to color--Black, White, Other--show the estimated percentages of each color group in the nation, each region, sex, parental education level and age and type of community classification. As shown in Exhibit B-1, the estimated percentage of 9-year-old Blacks (row 7) in the nation (column 1) is 13.6%, while of the total population of 9-year-olds in the Southeast (column 3), 24.8% are Black (row 7). The estimated percentage of 9-year-old Blacks living in the Southeast, 45.5%, is shown in row 1, column 4. The exhibits have been arranged so that for each reporting group identified by a column heading the percentages across the levels of a reporting variable identified by rows add to 100%. For example, the sum of the percentage of Blacks, Whites and Other equal 100% in every column of the exhibits.

[illegible]

[illegible]

[illegible]

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]

[illegible]

BEST COPY AVAILABLE

APPENDIX C

NATIONAL ASSESSMENT POLICY COMMITTEE

(Chairman of Policy Committee)
Mr. Robert McBride, President-Elect
Nat'l. Assoc. of School Boards of Education

Dr. George Brain
Dean, College of Education
Washington State University
Pullman, Washington

Mr. William R. Conway
Education Committee
Florida House of Representatives

Dr. Lyman V. Ginger
Superintendent of Public Instruction
State Department of Education
Frankfort, Kentucky

Dr. George Kozmetsky
Dean, Graduate School of Business
University of Texas

Ms. Joyce E. Lewis
Maine House of Representatives
Education Committee

Dr. Bill Lillard
Superintendent of Schools
Oklahoma City, Oklahoma

Dr. Frederick Mosteller
Chairman, Department of Statistics
Harvard University

Ms. Eleanor P. Sheppard
Chairman, Education Committee
Virginia House of Representatives

Dr. Stephen Wright
Vice President, College Entrance Examination Board

ex officio
Dr. Ralph W. Tyler
Science Research Associates
Chicago, Illinois

BEST COPY AVAILABLE

APPENDIX D

NATIONAL ASSESSMENT ANALYSIS ADVISORY COMMITTEE

Dr. Frederick Mosteller, Chairman since 1973. Professor of Mathematical Statistics, Harvard University (1970—).

Dr. John W. Tukey, Chairman 1965—73. Professor of Statistics, Princeton University (1965—).

Dr. Robert Abelson, Professor of Psychology, University of North Carolina (1965—72).

Dr. David Brillinger, Professor of Statistics, University of California at Berkeley (1973—).

Dr. William E. Coffman, Professor of Education, University of Iowa (1970—).

Dr. Lee J. Cronbach, Professor of Psychology and Education, Stanford University (1965—69).

Dr. James A. Davis, Director of National Opinion Research Center, Chicago, Illinois (1973—).

Dr. Janet Elashoff, Statistical Advisor, Center for Advanced Study in the Behavioral Sciences, Palo Alto, California (1973—).

Dr. John P. Gilbert, Staff Statistician at the Harvard University Computer Center (1970—).

Dr. Lyle V. Jones, Professor of Psychology, L.L. Thurstone Psychometric Laboratory, Dean of the Graduate School, University of North Carolina (1965—).

Dr. Lincoln Moses, Professor of Statistics Dean of the Graduate School, Stanford University (1974—).

Dr. Ralph W. Tyler, Science Research Associates, Chicago, Illinois (1965—69).